No-reference Stereoscopic Image Quality Assessment Using Natural Scene Statistics

Balasubramanyam Appina, Sameeulla Khan Md., and
Sumohana S. Channappayya

Lab for Video and Image Analysis (LFOVIA), Department of Electrical Engineering, Indian Institute of Technology Hyderabad, Yeddumailaram, India, 502205.
e-mail: {ee13m14p100001, ee13p1006, sumohana}@iith.ac.in.

Abstract

We present two contributions in this work: i) a bivariate generalized Gaussian distribution (BGGD) model for the joint distribution of luminance and disparity subband coefficients of natural stereoscopic scenes, and ii) a no-reference (NR) stereo image quality assessment algorithm based on the BGGD model. We first empirically show that a BGGD accurately models the joint distribution of luminance and disparity subband coefficients. We then show that the model parameters form good discriminatory features for NR quality assessment. Additionally, we rely on the previously established result that luminance and disparity subband coefficients of natural stereo scenes are correlated, and show that correlation also forms a good feature for NR quality assessment. These features are computed for both the left and right luminance-disparity pairs in the stereo image and consolidated into one feature vector per stereo pair. This feature set and the stereo pair's difference mean opinion score (DMOS) (labels) are used for supervised learning with a support vector machine (SVM). Support vector regression is used to estimate the perceptual quality of a test stereo image pair. The performance of the algorithm is evaluated over popular databases and shown to be competitive with the state-of-the-art no-reference quality assessment algorithms. Further, the strength of the proposed algorithm is demonstrated by its consistently good performance over both symmetric and asymmetric distortion types. Our algorithm is called Stereo QUality Evaluator (StereoQUE).

Keywords:
Natural scene statistics, stereoscopic images, no-reference image quality assessment.

1. INTRODUCTION

The volume of commercial and personal 3D/stereoscopic content being generated and consumed has dramatically increased over the past decade. This

trend could be attributed to a number of factors some of which include improved imaging and rendering devices [1] both at the industrial and consumer levels, increased commercial value for stereo content [2], the ever-increasing popularity of 3D gaming etc. Interestingly, the statistics gathered by the Motion Picture Association of America (MPAA) [3] assigns an exclusive category to 3D (stereo) movies. A variety of parameters such as viewer age, attendance, box office numbers, 3D screens count are tracked in these reports. According to these statistics, 3D movies have seen a consistent increase in revenue over the past decade and are poised to continue the trend moving forward. A stand-out number in this report is that 8 of the top 10 highest grossing movies in 2013 were released in 3D!

Given this upward trend, the role of a good objective measure of perceptual stereoscopic quality cannot be overemphasized. A good objective quality measure can not only serve as a proxy for subjective evaluation but can also be used as an objective function for the optimal design of stereoscopic image processing systems. A natural or straightforward approach to stereoscopic quality assessment would be to simply use existing 2D image quality metrics by applying them to the left and right views of a stereo pair and averaging the scores. While such an approach appears reasonable, it is not clear that it would be successful. As noted by several authors [4, 5, 6, 7], the effects of distortion (especially asymmetric distortion) on depth perception makes stereo quality assessment more challenging than a simple extension of 2D approaches.

As with 2D image quality assessment (IQA) methods, stereoscopic image quality assessment algorithms are also classified into three categories depending on the usage of the pristine stereo image pair for quality assessment. Full reference (FR) algorithms make use of the reference stereo pair in their entirety, while reduced reference (RR) approaches make use of partial reference information for quality evaluation of the test stereo pair. No reference (NR) methods on the other hand, do not use reference information in evaluating the perceptual quality of a stereo image pair.

In this paper, we present a no-reference algorithm for the assessment of stereoscopic image quality. Our approach is inspired by the success of applying natural scene statistics (NSS) to all the classes of 2D IQA algorithms – FR, RR [8], and NR [9, 10, 11, 12] methods. A recent statistically motivated approach for NR stereo quality assessment by Su et al. [13] further inspired our work. We first propose a bivariate generalized Gaussian density (BGGD) model [14] for the joint statistics of luminance and disparity subband statistics of natural stereo scenes. We then use this model in the design of a NR stereo IQA algorithm we call Stereo QUality Evaluator (StereoQUE).

Our paper is organized as follows. We review relevant literature in Section 2, followed by a description of our algorithm in Section 3. We first describe the proposed statistical model in Section 3.1 and then present the proposed NR stereo IQA algorithm in Section 3.2. We discuss our results in Section 4 and present concluding remarks in Section 5.

2. BACKGROUND

Stereoscopic image quality assessment in all its flavors (FR, RR and NR) has received relatively less attention compared to 2D image quality assessment. This could be attributed to several factors including but not limited to the smaller fraction of stereo content compared to 2D content, an inclination to apply 2D metrics to each of the left and right stereo images, and paradoxically, the visual annoyance resulting from poorly designed stereo imaging systems. NR stereo quality assessment has received even lesser attention than FR and RR methods. We first review significant human visual system (HVS) inspired FR, RR and NR stereo image quality assessment algorithms. To place our proposed algorithm in perspective, we briefly review NSS-based 2D and stereo NR-IQA algorithms, followed by a review of natural stereoscopic scene statistical models.

2.1. HVS-inspired Stereoscopic IQA Algorithms

Campisi et al. [15] were among the first researchers to carry out a systematic study of stereoscopic image quality assessment. Their study involved both subjective and FR objective quality evaluation. Objective evaluation was carried out by applying existing 2D FR-IQA algorithms separately to the left and right image reference-test pairs. Perceptually inspired approaches were used to combine the left and right scores. They concluded that 2D FR-IQA algorithms do not necessarily perform well on stereo images, and that further investigation was required. Stereo Band Limited Contrast (SBLC) [16] is a HVS-inspired FR stereo IQA algorithm where the ratio of contrast to relative luminance at edge and corner image locations was found to give a good perceptual quality estimate. Neither of these early FR methods explicitly used depth information.

Benoit et al. [17] proposed an important FR algorithm that explicitly considered disparity information in its analysis. They applied the 2D FR-IQA algorithms to the spatial information of stereoscopic reference and distorted images. In their approach, they considered the disparity information as an extra feature and included it both locally and globally in their algorithm. The performance of this algorithm was shown to be competitive over commonly occurring distortions [18]. You et al. [19] proposed a similar methodology to measure local and global similarity of luminance and disparity images that performs well on both a custom database and the LIVE database [18]. Chen et al. [20] studied the effect of binocular depth variation on the perceptual quality of stereoscopic images. They showed that the overall quality of experience decreases with increasing depth. Further, they showed that for 2D images, the quality of experience (QoE) is not affected by depth variation. Didyk et al. [7] studied the effect of compression artifacts on the binocular disparity estimation. They concluded that depth perception in a stereoscopic image decreases with increase in compression artifacts. Wang et al. [4] conducted a systematic study on the effects of symmetric and asymmetric distortions on stereo images. They found that for symmetrically distorted images, the overall quality score of an image is equal to the average of the left and right views. However, they found that simple averaging did not work in the case of asymmetric distortion of views.

Bensalma and Larabi proposed a series of HVS-inspired FR approaches based on binocular energy modeling [21, 22, 23]. They rely on explicitly modeling simple and complex cells in the visual cortex as a series of multiscale oriented band pass filters (implemented using a combination of complex wavelet and bandelet transforms). The difference in the subband energies of the reference and distorted images is used to compute a perceptual quality score. Though they explored the binocular energy modeling, they did not consider the effect of binocular suppression [24] on perception.

Ryu et al. [25] proposed a FR-IQA metric for stereoscopic images based on the perceptual binocular model and structural similarity (SSIM) parameters (i.e. luminance, contrast, structural similarity). These parameters are computed for each image of the stereo pair and pooled together with different weighting combinations to come up with a single quality score. Depth or disparity information was not used in this work. Hachicha et al. [26] proposed a FRQA metric for stereoscopic images based on modeling of HVS and its properties. They considered the concepts of contrast masking effect and Binocular Just Noticeable Difference (BJND) to propose metric and performed their evaluation on the LIVE phase I dataset. Zhang et al. [27] proposed a FR stereoscopic IQA metric based on 3D MAD estimation algorithm. To proposed this algorithm, they applied the conventional MAD estimation on left and right images of stereoscopic view and they combined these scores with the block based contrast measured scores. Li et al. [28] proposed a stereoscopic FR-IQA metric based on structure and texture decomposition but without taking depth/disparity information into account. They measured gradient magnitude similarity, luminance and contrast similarity from the structure and texture decompositions. They compute these values for each view and finally combine these scores to come up with a single score for a stereo pair. Shen et al. [29] proposed a stereo FR-IQA metric based on a segmentation of the disparity map using SSIM values. First, they compute the disparity maps for original and distorted stereo pairs. These disparity maps are then segmented, and the SSIM scores are computed on the segmented maps. Lin et al. [30] proposed a setereoscopic FR-IQA metric based on a combination of 2D metrics and binocular frequency integration. In addition, they evaluated their method on synthesized color-plus-depth 3D images. Hu et al. [31] proposed a FR-IQA method based on a distortion separation method. They compare reference and distorted images individually i.e., left and right images separately but again did not use depth information. To find the structural similarity between these images, they rely on singular value decomposition (SVD) and phase-amplitude description (PAD) methods. These scores are pooled to compute a quality score for the stereo image pair. C et al. [32] proposed a stereoscopic full reference IQA metric based on modeling the mechanisms of HVS 3D perception. They used the modeling of simple and complex cells based on the behavior of binocular supression theory. Cardoso et al. [33] proposed a stereoscopic FR & NR IQA metric based on disparity weighting method. They showed the effect of utilizing the disparity information in stereoscopic metrics. Zhou et al. [34] proposed a RR-IQA metric based on the degradation level of water marking on symmetric distorted images. They embedded water mark-

ing on the extracted features and estimate distortion based on the amount of degradation experienced by the watermark.

The cyclopean image paradigm has proven to be a successful approach for integrating depth information into a 2D image and subsequently relying on 2D IQA algorithms for quality assessment. Maalouf and Larabi [35] proposed a RR-IQA stereoscopic metric based on the cyclopean model of stereoscopic image and color disparity maps. They construct individual cyclopean images and color disparity maps for the reference and distorted set and find the HVS-based thresholds. Later, they compare the sensitivity coefficients of reference and distorted images to predict the quality score. Chetouani [36] proposed a stereoscopic FR-IQA metric that is also based on the cyclopean image paradigm and 2D image quality metric (IQM) fusion. In their approach they applied the 2D IQA metrics on the cyclopean image. They extracted 2D features from the cyclopean image and used an artificial neural network to predict the quality score. Chen et al. [18] proposed a FR-IQA metric for stereoscopic images based on the cyclopean paradigm. They considered the effect of binocular rivalry on stereoscopic images that are asymmetrically distorted. They showed the effect of disparity information usage on the overall quality score, and concluded that disparity information is indeed useful.

Akhter and Sazzad [37, 38] present a NR stereo IQA algorithm inspired by the HVS. Local edge-based and non-edge based disparity estimates, along with blur and blockiness estimates are used as distortion discriminating features. However, the effect of binocular rivalry was not explicitly considered in their work. The authors demonstrate the efficacy of their method on a database comprised exclusively of JPEG artifacts. Ryu et al. [39] proposed a stereoscopic NR-IQA algorithm based on high and low frequency components of left and right luminance images. They extracted the blurriness, saliency, blockiness map from each stereoscopic image. They model these maps and used as quality predictor of the stereoscopic image. Shao et al. [40] proposed a NR-IQA metric based on distortion specific features of luminance information. However, they did not include disparity information in their metric. They used these features to characterize the type of the distortion. Support vector regression (SVR) is used to estimate the quality of a test stereo pair from its features. Fezza et al. [41] proposed a NR-IQA metric for stereoscopic images based on local blurriness maps and local entropy values computed with the help of a threshold. Two blurriness maps are computed for each stereo pair of images and a threshold value is chosen based on the occluded regions to compute the blurriness scores. The left and right image scores are pooled with the help of local spatial entropy values. The average of the occluded and non-occluded region scores is assigned to be the quality of the stereo image pair. Shao et al. [42] proposed a NRQA metric for stereoscopic images using a binocular guided quality lookup and visual codebook. Their metric involves the construction of phase-tuned quality lookup (PTQL) and phase-tuned visual codebook (PTVC). They find the gradient similarities and centroid clustering of binocular energy responses to construct PTQL and PTVC. Gu et al. [43][44] proposed a stereoscopic NRQA metric based on ocular dominance theory and parallax effect between left and

right views of a stereoscopic pair. The ocular dominance was calculated based on the distress function and degree of parallax was calculated based on central region of left and right images of the stereoscopic view. Jiang et al. [45] proposed a stereo NR-IQA method based on the cyclopean model and used a 2D saliency model to predict the perceptual quality. Solh et al. [46] proposed a stereoscopic NR video quality metric based on temporal outliers (TO), temporal inconsistencies(TI), spatial outliers(SO). They used depth based rendering (DIBR) to generate stereoscopic videos. Further, they used the TO, TI and SO to evaluate the ideal depth estimation.

2.2. NSS-inspired NR-IQA Algorithms

While deterministic (HVS-inspired) approaches to FR, NR, and RR IQA have been quite popular and successful, statistical approaches have also proven to be very successful. Our proposed algorithm is motivated by the success of these statistical approaches. We now review a small but significant subset of NSS-inspired 2D and stereo NR-IQA methods.

The generalized Gaussian density (GGD) and the Gaussian scale mixture (GSM) density are two models for 2D natural scene statistics [47, 48] that are very popular and have been widely employed in 2D IQA. Moorthy et al. [49] studied NSS feature (GSM and GGD model parameter) variations over natural and unnatural scenes. They proposed a NR-IQA metric for 2D images based on a two-stage frame work. With the help of NSS features they classified distorted image into different distortion classes and used these statistical features to predict the quality score of the distorted image. Mittal et al. [50, 51] proposed a 2D NR-IQA metric based on statistical modeling of natural scenes in the pixel domain. They considered mean subtracted contrast normalized (MSCN) luminance coefficients in the spatial domain and computed pairwise product of adjacent normalized luminance coefficients to get the distortion specific information. Asymmetric GGD fitting model was used to fit the MSCN coefficients in their method to discriminate the type of distortion. Saad et al. [11] proposed a 2D image NR-IQA metric based on NSS modeling of DCT coefficients, again using a univariate GGD model. Their method does not use any distortion specific features. They applied the sample Bayesian inference model to predict the quality score of an image. Sheikh et al. [52] proposed a 2D image NR-IQA metric based on the NSS modeling of JPEG compressed images. They showed that the deviation or variation of expected natural statistics of respective image is related to the perceptual quality. Mittal et al. [53] proposed a NR stereoscopic quality of experience assessment metric based on the statistics of luminance images and disparity maps. A slew of statistics including mean, variance, skewness, kurtosis, differential disparity etc. are computed on the disparity maps. The mean, variance and kurtosis of spatial activity maps of the left and right images are also computed. These statistics compose the feature vector for the stereo pair. Dimensionality reduction techniques such as principal components analysis (PCA) and forward feature selection (FFS)are employed on a training images to reduce feature vector dimension. Finally, linear regression is used to estimate the perceived quality of experience. For video quality assessment, the

same method is applied on a frame-by-frame basis. However, it is important to note that this approach, though based on statistical features, does not attempt to model and exploit the rich and unique signature of natural stereo scene statistics. In the following, we review statistical models of natural stereoscopic scenes.

2.3. Stereoscopic NSS Models

The statistical characteristics of natural stereoscopic scenes have been studied by a number of researchers. For stereoscopic images, the modeling is typically done on subband luminance and range/disparity coefficients – either jointly or marginally. Huang et al. [54] model the range statistics of the natural environment. To construct these statistics they used a laser range finder [55] to acquire the range maps of the natural environment. To describe the properties of range maps they consider the single-pixel and derivative statistics. The discontinuities of range maps are explored using Haar wavelets. Liu et al. [56] construct the disparity information from the depth map which is created by the range finder with a fixed point described. To construct these maps, a spherical model of the eye structure is employed. Finally, they correlate the disparity distribution with the stereopsis of the HVS.

Potetz and Lee [57] and Liu et al. [58, 59] explored the statistical relationship between the luminance and disparity maps in the spatial domain and in a multiscale subband domain respectively. They modeled conditional histograms of luminance and disparity/range using a univariate GGD [60, 61] model. They also demonstrated considerable correlation between the luminance and range/disparity subband coefficients. Khan et al. [62] proposed a FR-IQA metric for stereoscopic images based on the statistical modeling of luminance and disparity. They modeled marginal statistics of luminance and disparity using a univariate GGD model. In their approach, they considered the disparity maps only for reference stereoscopic images. Chen et al. [63] also proposed a NR-IQA metric for stereoscopic images based on the cyclopean model and NSS features. In this method, they again considered the effect of binocular rivalry on perceptual quality. In their approach, they used a combination of 2D features extracted from the cyclopean image and 3D features extracted from the uncertainty map produced by the stereo matching algorithm.

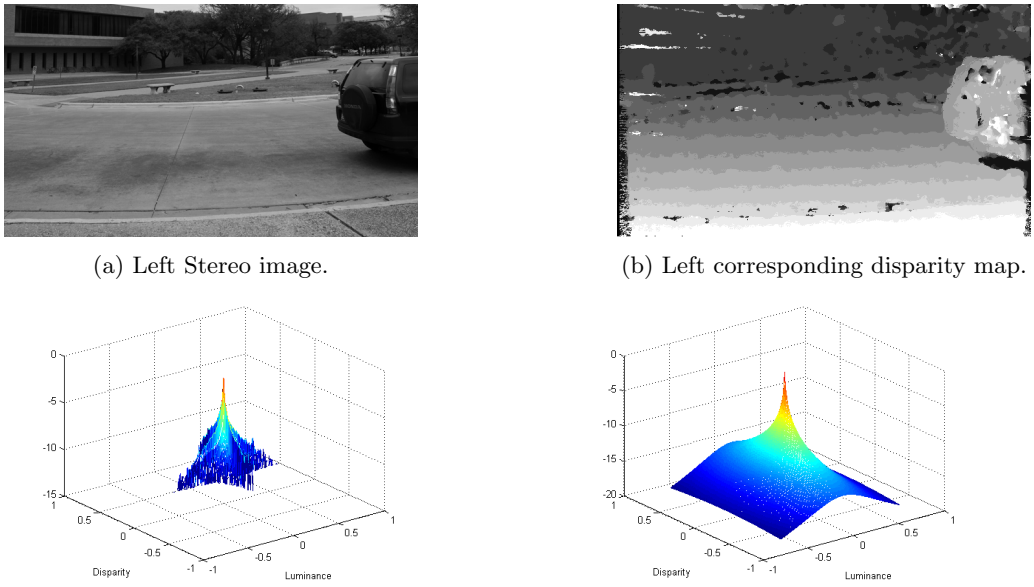
Su et al. [64] conducted a study on the relationship between the chrominance and range components. They showed that the distribution of the range gradients conditioned on the Gabor responses of chrominance images can be modeled with a Weibull distribution. Su et al. [13] proposed a stereoscopic NR-IQA metric based on the bivariate statistical correlation model [14] (a bivariate GGD model) for spatially adjacent luminance subband coefficients. This model is applied to capture the joint statistics of subband coefficients of a converged cyclopean image (CCI). In addition to the bivariate statistics, they also rely on univariate NSS features extracted from the CCI to form their feature vector. Support vector machine training is performed using these features and DMOS scores as labels. Support vector regression is used to predict the quality of a test stereo image pair. This algorithm is the current state-of-the-art NR stereo IQA

technique and outperforms state-of-the-art FR stereo IQA algorithms as well. We present a detailed comparison of the proposed algorithm with the methods by Chen et al. [63] and Su et al. [13] in Section 4.

3. PROPOSED ALGORITHM

Our proposed algorithm is closest in philosophy to the work by Su et al. [13]. As noted previously, an important feature in their method is the BGGD modeling of adjacent luminance subband coefficients of the CCI. In this section and in Section 4, we describe the proposed methodology and demonstrate how it is different from previous NSS-inspired NR stereo IQA algorithms.

Earlier works by Potetz and Lee [57] and Liu et al. [58] only consider the marginal statistics of luminance and disparity maps or their subband coefficients. We strongly believe that the joint statistics capture crucial information about the relation between luminance and depth/disparity subband coefficients. As we show in the following sections, this information could be useful in quantifying the effects of distortion on perceptual quality. We propose a model to capture the joint luminance and disparity/depth subband statistics of natural stereoscopic scenes. We empirically show that a bivariate GGD is an excellent model for the joint statistics. We present this model first followed by a description of the proposed NR stereo IQA algorithm dubbed Stereo QUality Evaluator (StereoQUE).



(a) Left Stereo image.

(b) Left corresponding disparity map.

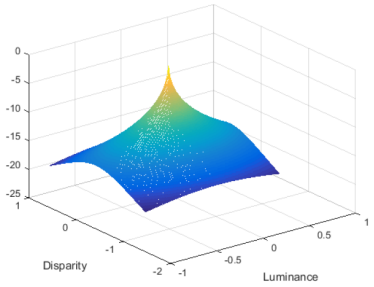
(c) Joint histogram.

(d) Bivariate GGD fit.

Figure 1: BGGD modeling. $\alpha = 9 \times 10^{-16}, \beta = 0.11, \chi = 5 \times 10^{-8}$.

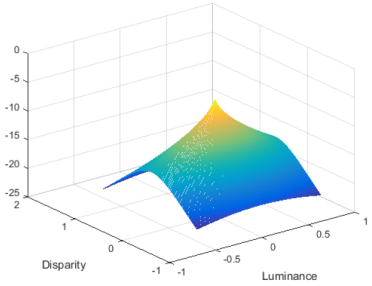
At the outset we would like to note that our analysis is carried out on a multi-scale/multi-orientation subband decomposition of the stereo images and their disparity maps. This is motivated by the multiscale modeling of the human visual system [58]. Specifically, we work with a steerable pyramid decomposition [65, 66] due to the advantages of translation-invariance and rotation-invariance compared to wavelet decompositions [67] [68]. In our work we used 3 spatial scales and six orientations $(0°, 30°, 60°, 90°, 120°, 150°)$ in the decomposition. The subband decomposition is performed on the logarithm of the luminance and disparity images. The logarithm of the images is taken to mimic the visual system as demonstrated by several authors [69] [58] [57].

(a) Reference left image.



(b) BGGD fit for reference image.

(c) Distorted left image.



(d) BGGD fit for distorted image.

Figure 2: Illustration of BGGD variation over reference and distorted image.

3.1. Bivariate GGD Modeling

The multivariate GGD distribution of a random vector $\mathbf{x} \in \mathbb{R}^N$ is given by [14]

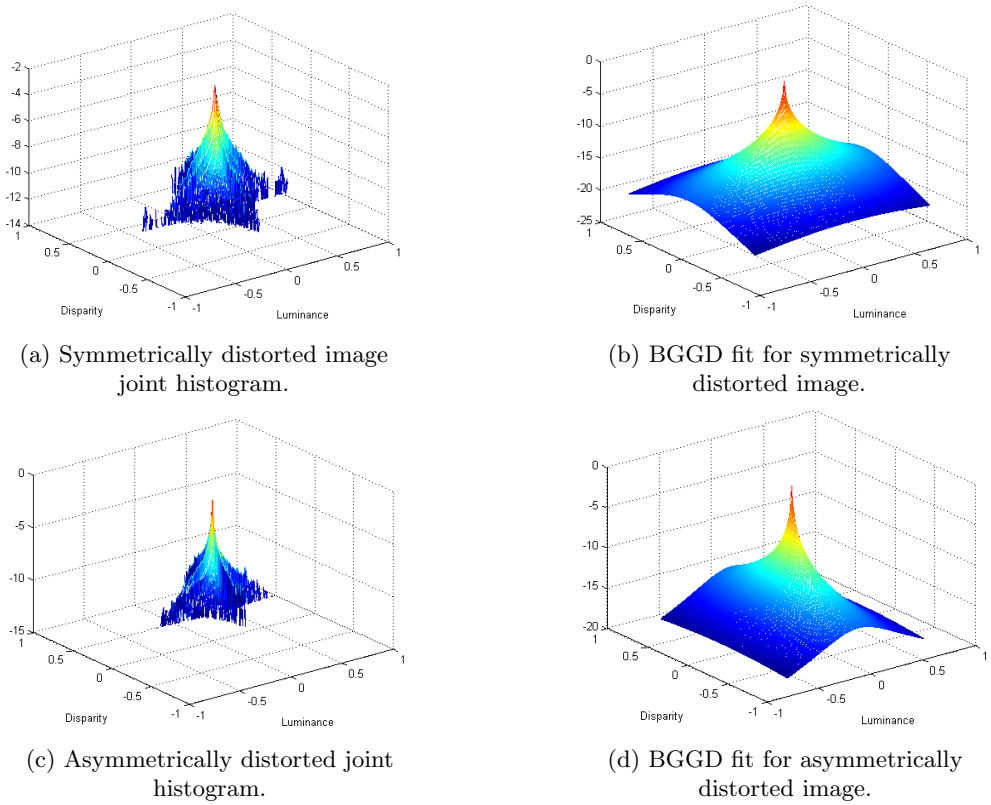
$$p(\mathbf{x}|\mathbf{M}, \alpha, \beta) = \frac{1}{|\mathbf{M}|^{\frac{1}{2}}} g_{\alpha,\beta}(\mathbf{x}^T \mathbf{M}^{-1} \mathbf{x}), \tag{1}$$

$$g_{\alpha,\beta}(y) = \frac{\beta \Gamma(\frac{N}{2})}{(2^{\frac{1}{\beta}} \Pi \alpha)^{\frac{N}{2}} \Gamma(\frac{N}{2})} e^{-\frac{1}{2}(\frac{y}{\alpha})^{\beta}}, \tag{2}$$

where $\mathbf{M}$ is an $N \times N$ symmetric scatter matrix, α is the scale parameter, β the shape parameter and $g_{\alpha,\beta}(\cdot)$ is the density generator. Its heavy-tailed

and completely parameterized form makes it particularly attractive in modeling natural scenes.

Fig. 1a shows the left image of a pristine stereo pair chosen from the LIVE Phase-II [18] database. Fig. 1b show the disparity map computed using an SSIM-based algorithm [18]. Fig. 1c shows the joint histogram between the log luminance and log disparity subband coefficients at the first scale and 0° orientation of a steerable pyramid decomposition. From the joint histogram we can observe a clear sharp peak and heavy tails. We have observed a similar trend across a wide range of images across various stereo image databases at various scales and orientations. Based on these observations, we propose that a bivariate GGD ($N = 2$) can accurately model the *joint* distribution of luminance and disparity subband coefficients.



(a) Symmetrically distorted image joint histogram.

(b) BGGD fit for symmetrically distorted image.

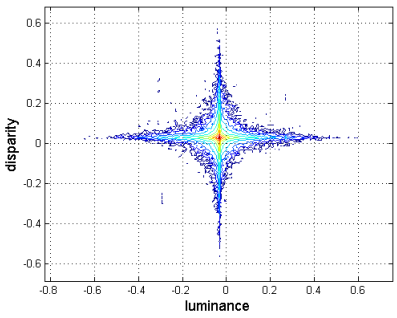
(c) Asymmetrically distorted joint histogram.

(d) BGGD fit for asymmetrically distorted image.

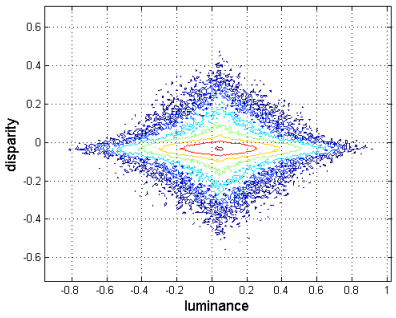
Figure 3: Illustration of BGGD variation over symmetric and asymmetric distortions.

In our evaluation of the proposed BGGD model over the LIVE Phase-I, Phase-II, and the MICT [38] databases, the highest value for χ (indicating the worst fit) was found to be of the order of 10^{-7}. The low values of χ present excellent statistical support for our hypothesis. We also found that the shape

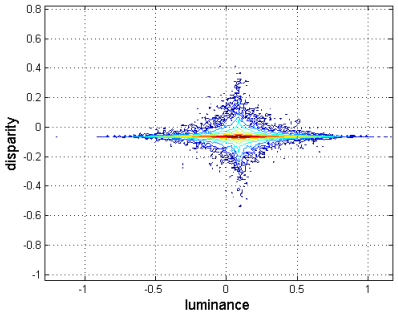
parameter β lies in the interval [0.0014, 0.45], corresponding to heavy tailed distributions ($\beta = 1$ corresponds to the Gaussian distribution).



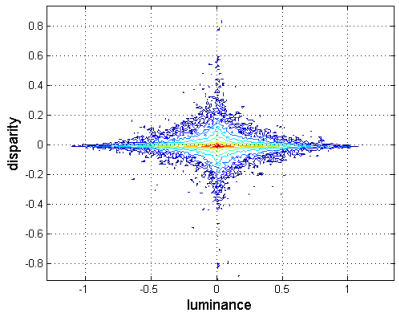
(a) Reference
($\alpha = 9.3 \times 10^{-16}, \beta = 0.11$).



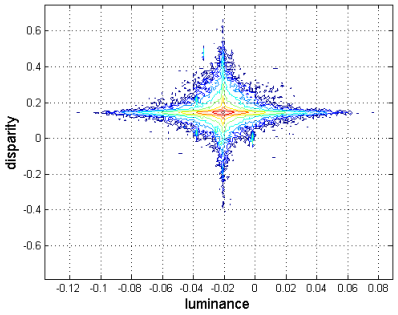
(b) AWGN
($\alpha = 1.1 \times 10^{-4}, \beta = 0.34$).



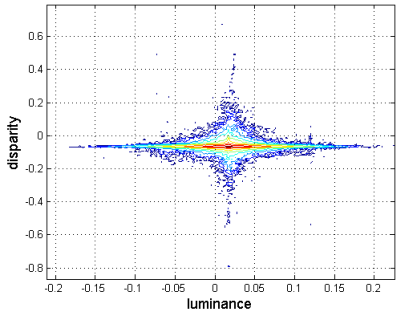
(c) JPEG2000
($\alpha = 1.38 \times 10^{-10}, \beta = 0.15$).



(d) JPEG
($\alpha = 4 \times 10^{-19}, \beta = 0.01$).



(e) Gaussian blur
($\alpha = 1.5 \times 10^{-8}, \beta = 0.21$) .



(f) Fast fading
($\alpha = 2.23 \times 10^{-8}, \beta = 0.2$).

Figure 4: Joint histogram contours for commonly occurring distortions.

This observation provides corroborative evidence for the findings of Liu et al. [58], and clearly highlights the non-Gaussian nature of the joint statistics.

We now present several examples to demonstrate that the proposed BGGD model is effective even in the presence of distortions. Figs. 2a and 2b shows an undistorted image and its BGGD model fit. Fig. 2c shows a distorted version of 2a where white Gaussian noise (AWGN) been added. Due to AWGN, the tails are heavier and the peak is shorter. We see from Fig. 2d that the BGGD fit is able to capture this effect. Fig. 3 shows the variation of the joint statistics of the same stereo image (as in Fig. 1) when subjected to symmetric and asymmetric distortion. These are also shown for the first scale and 0° orientation.

We now present the effect of commonly occurring distortions on the joint distribution of luminance and disparity subband coefficients. Fig. 4 shows the contour plots of the joint histogram of an undistorted image and several distorted versions of it (from the LIVE Phase-II database). These histograms have been computed at the first scale and 0° orientation subband. From the plots, it is clear that distortions affect the joint distribution and therefore the best-fitting BGGD model parameters. Further, the change in the model parameter values relative to the reference is proportional to the amount of distortion. This observation is fundamental to our premise that the BGGD model parameters form good discriminatory features for the quality assessment task. In addition to the BGGD model parameters, we rely on the result by Liu et al. [58] that there exists significant correlation between luminance and disparity subband coefficients.

Fig. 5 shows the variation of correlation across the six orientations at the first scale for the same set of reference and distorted images used for the joint histogram computation. As with the joint histogram, Fig. 5 shows that correlation is also affected due to distortion in a manner that is proportional to the amount of distortion. Therefore, correlation also serves as a good feature for no-reference quality assessment. Our proposed NR stereo IQA algorithm is built on this NSS-inspired foundation and is presented in the following.

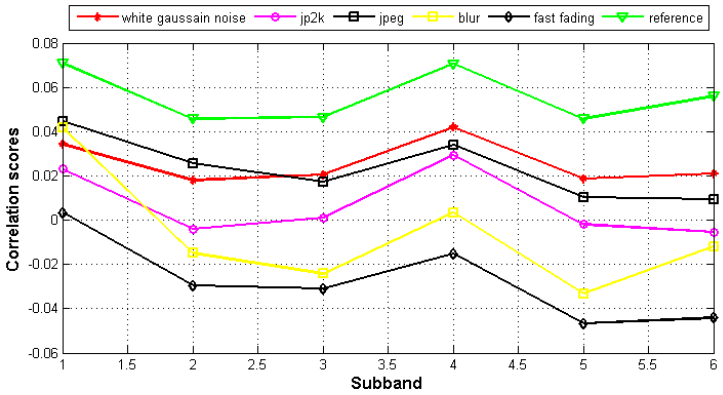


Figure 5: Correlation scores for LIVE Phase-II database.

3.2. Stereo QUality Estimator (StereoQUE) Algorithm

The flowchart of the proposed algorithm is shown in Fig. 6 and is described in the following subsections.

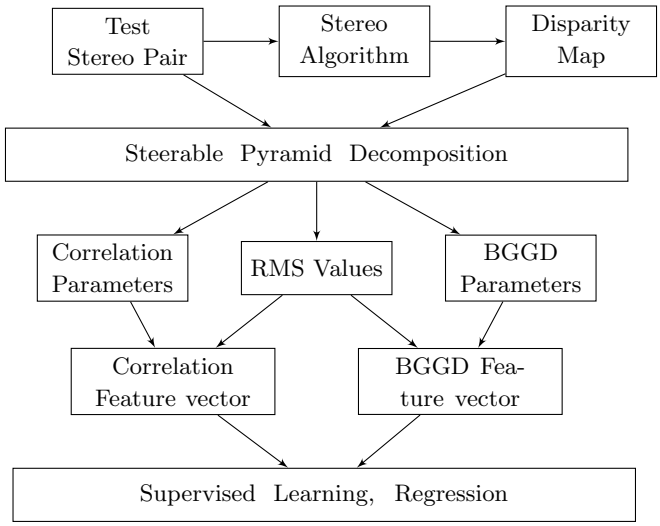


Figure 6: Flowchart of proposed algorithm.

3.3. Feature Extraction

Given a stereo image pair, disparity maps are first estimated using an SSIM-based disparity algorithm [18] and we limited the pixel variation in disparity map upto 15 pixels. We chose the SSIM based stereo matching algorithm since it provides a good trade-off between accuracy and time complexity [70]. Two sets of disparity maps are computed per stereo pair – one with the left image treated as reference and the other with the right image treated as reference. As mentioned in Section 3, our analysis is carried out on oriented subbands of the luminance and disparity images. Specifically, we operate at 3 scales and 6 orientations resulting in 18 subbands. At each subband, the BGGD model parameters (α, β) and the correlation coefficient γ are computed. The features corresponding to the left and right images are subscripted by l and r respectively. This results in two feature vectors $\mathbf{f}_l = [\alpha_l^1 \dots \alpha_l^{18}; \beta_l^1 \dots \beta_l^{18}; \gamma_l^1 \dots \gamma_l^{18}]^T$, $\mathbf{f}_r = [\alpha_r^1 \dots \alpha_r^{18}; \beta_r^1 \dots \beta_r^{18}; \gamma_r^1 \dots \gamma_r^{18}]^T$, where the superscript corresponds to the subband index. Thus we end up with a pair of feature vectors $(\mathbf{f}_l, \mathbf{f}_r)$ containing a total of 108 elements per stereo image pair.

3.4. Feature Consolidation

Once the feature vectors are extracted, the next step in our algorithm is supervised learning using DMOS scores as labels. While a supervised learning

algorithm could be trained with the feature vector pair $(\mathbf{f}_l, \mathbf{f}_r)$, it is worthwhile exploring a reduction of the feature space dimensionality. Dimensionality reduction is further motivated by the fact that stereo images typically contain redundant information, and by implication, so would the pair $(\mathbf{f}_l, \mathbf{f}_r)$. One way to reduce dimensionality would be to consolidate the vector pair into a single vector. To achieve this goal, we propose the following consolidation strategy that is motivated by the fact that binocular strength is a convex sum of monocular strengths [71]. The convex weights depends on relative eye dominance that is particularly visible in the case of asymmetric distortion. Khan et al. [62] considered that stimulus strength directly relates with band strength of each subband and use root-mean-square (RMS) of subband of as its band strength. Also, [62] considered only band strengths of luminance subbands because they relied only on luminance statistics. Here, we find statistical parameters from both luminance and disparity subbands. Therefore, to effectively consolidate the feature vectors we consider the RMS of both luminance and disparity subbands. The convex weights are obtained as follows:

$$w_l^i = \frac{R_l^i R_{dl}^i}{R_l^i R_{dl}^i + R_r^i R_{dr}^i}, \tag{3}$$

$$w_r^i = \frac{R_r^i R_{dr}^i}{R_l^i R_{dl}^i + R_r^i R_{dr}^i}, \tag{4}$$

$$\mathbf{f} = \mathbf{w}_l \odot \mathbf{f}_l + \mathbf{w}_r \odot \mathbf{f}_r, \tag{5}$$

where w_l^i, w_r^i are the weights assigned to the i^{th} left and right subbands respectively, R_l^i, R_r^i are the RMS values of the i^{th} left and right luminance subbands respectively, R_{dl}^i, R_{dr}^i are the RMS values of the i^{th} left and right disparity subbands respectively. The vectors $\mathbf{w}_l = [w_l^1, \ldots, w_l^{18}]^T, \mathbf{w}_r = [w_r^1, \ldots, w_r^{18}]^T$ and $\odot$ denotes element-wise product. The weights are chosen to be proportional to the stimulus strength to the respective eye and is inspired by binocular rivalry where the dominant stimulus strongly influences overall perception. The RMS values of the luminance and disparity subbands are chosen to be representatives of the respective subband strengths. The final feature vector $\mathbf{f}$ is a convex combination of $\mathbf{f}_l$ and $\mathbf{f}_r$. It is worthwhile recalling that the cyclopean image is also a convex combination of the left and right luminance images of a stereo pair.

3.5. Supervised Learning and Regression

The final stage of our algorithm is supervised learning using the feature vector $\mathbf{f}$ of a stereo pair (constructed using (5)) and its DMOS score as the feature label. We use a support vector machine for regression (SVR) estimating the score of a test input (after training with labeled data). SVMs can offer good performance even when the number of training points available are small in addition to advantages like being more accurate in one-versus-rest schemes [72], finding the global minimum [73], sparseness of the solution [74] etc. In

our experiments with kernel choices, a radial basis function (RBF) kernel was found to give the best performance. We report our results for this kernel choice. For each database, 80% of the images were used to train the SVM and the rest were used for regression. The open-source SVM package *LIBSVM* [75], citewebsite:libsvmpack was used in our experiments. For statistical consistency, we perform the training and testing 5000 times. For each run, we perform a random image assignment to the train and test sets and also ensure that the train and test sets do not overlap. We report results that are averaged over these 5000 trials.

4. RESULTS AND DISCUSSION

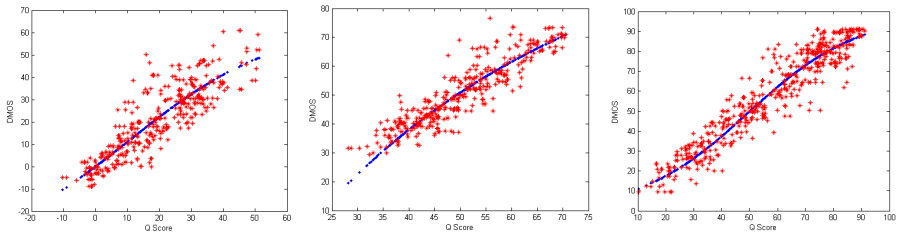
The performance of our algorithm was evaluated on the LIVE Phase-I [76], Phase-II [18, 63] and the MICT [38] stereo image databases. Both the LIVE databases consist of five distortion categories: Additive White Gaussian Noise (AWGN), JPEG2000 (JP2K), JPEG, Gaussian Blur (BLUR), and Fast Fading (FF). LIVE Phase-I contains 20 pristine stereo pairs and 365 distorted stereo pairs. Each distortion type has 4 levels of distortion strengths. LIVE Phase-II contains 8 reference stereo pairs and 360 distorted stereo pairs. Each distortion type has 9 levels of distortion strengths. The LIVE Phase-II database are also classified into symmetric (120 stereo pairs) and asymmetric (240 stereo pairs) distortions. The MICT database is composed of 13 reference stereo pairs and JPEG distorted versions. The database contains 546 asymmetrically distorted stereo pairs and 78 symmetrically distorted ones. The DMOS scores reported in the databases were used in the training process and for the performance evaluation of the algorithm on the test stereo pairs. We would like to note that in the case of the MICT database, only MOS scores have been published. The DMOS scores for the distorted images were computed using the appropriate reference stereo pair MOS values. A performance comparison of StereoQUE with the state-of-the-art FR, RR and NR stereo IQA algorithms is presented in the following tables.

Following the standard procedure suggested by the Video Quality Experts Group (VQEG) [77], we applied a nonlinear regression with a 4-parameter logistic transform to the output of the algorithm before evaluating its performance. Specifically, we used the logistic transform recommended by VQEG

$$f(x) = \frac{\tau_1 - \tau_2}{1 + \exp(\frac{x-\tau_3}{|\tau_4|})} + \tau_2,$$

where x denotes the raw objective score, and τ_1, τ_2, τ_3 and τ_4 are the free parameters selected to provide the best fit of the predicted scores to the DMOS values. Fig. 7 shows the scatter plots for LIVE phase-I, LIVE phase-II and MICT databases.

Tables 1, 2, 4, 5 and 8 show the LCC and SROCC scores of StereoQUE on LIVE Phase-I, LIVE Phase-II and MICT database. The results reported on the MICT dataset are over the same set of images as in Chen et al. [18]. It is clear



(a) LIVE PHASE-I. (b) LIVE PHASE-II. (c) MICT.

Figure 7: Scatter plots of StereoQUE versus DMOS for various stereo image databases.

that the proposed metric's scores are able to mimic the human subjective scores well. Also, we highlight the performance evaluation on asymmetric distortions in Table 7. Additionally, we compared the results of LCC and SROCC with different state-of-the-art stereoscopic IQA metrics. The FR and RR methods are in normal font while the NR methods are *italicised*. From these results it is clear that StereoQUE performs competitively not just with NR algorithms but also with FR and RR algorithms. The tables 3 and 6 shows the performance results of RMSE on LIVE Phase-I and LIVE Phase-II dataset. It is clear that, the proposed metric StereoQUE shows good performance with DMOS scores. The consistent performance of StereoQUE as measured using LCC, SROCC, and RMSE show that is possesses good linear correlation with DMOS, good ranking based monotonicity with DMOS, and finally good accuracy relative to DMOS and thereby showcasing its overall strength.

We now present a qualitative discussion of our results. Firstly, we explain the performance of our algorithm vis-a-vis the choice of our NSS features. Tables 1 - 8 show that StereoQUE performs competitively with the state-of-the-art methods over all the databases considered. The performance of StereoQUE is consistently high on asymmetric distortions and does better than the state-of-the-art FR, RR, and NR stereo IQA methods. As shown in Figs. 1, 2 the proposed BGGD model is able to accurately model the joint statistics of the luminance and disparity subband coefficients. Importantly, as shown in Fig. 4, variation in the model parameters is proportional to the amount of distortion in the stereo pairs, thereby making it effective for measuring stereo image quality. Further, the state-of-the-art performance over asymmetric distortions can be explained by our feature consolidation strategy that effectively weights the left and right views in proportion to their overall contribution to distortion perception.

We now compare our approach with the two state-of-the-art NR stereo IQA algorithms proposed by Chen et al. [63] and Su et al. [13] in terms of the underlying principles, and computational complexity. At a high level, all the approaches rely on NSS-based features that are used to train a support vector machine (SVM). The objective score of a test stereo pair is estimated using support vector regression (SVR). The difference in the approaches is in the

details that we bring out in the following. Both Chen et al. and Su et al. extract 2D and 3D NSS features from the stereo pair. They rely on the cyclopean paradigm for 2D NSS feature extraction. The 2D features are the GGD model parameters of the MSCN coefficients of the cyclopean image. This is inspired by the work of Mittal et al. [50]. Su et al. extract GGD features from the steerable pyramid decomposition of the cyclopean image as well. The two methods differ in their approach for extracting 3D NSS features. Chen et al. [63] model the locally-normalized disparity maps using a GGD and use the model parameters as features. Additionally, they also apply a log-normal model to the SSIM map computed between the left view and disparity compensated right view. This SSIM map is termed an uncertainty map. The parameters of this GGD model also constitute their 3D features. Su et al. [13] extract 3D features from the parameters of a BGGD model of luminance wavelet coefficients and from the parameters of an exponential model used to capture the correlation between subband coefficients. Our method does not use 2D features explicitly. Instead, its features are extracted from the parameters of the proposed BGGD model of the joint histogram of the luminance and disparity subband coefficients, and the correlation between these coefficients.

The choice of our features results in computational complexity that is comparable to the two state-of-the-art methods. We present a qualitative comparison since the code for neither of the two state-of-the-art methods is freely available. We compare the computational complexity of the multiscale decomposition and feature extraction stages since these are most expensive stages in all the algorithms. We do not include the complexity involved in finding disparity maps or support vector training and regression since these stages are common to all the methods.

In Chen et al. [63], the most computationally expensive stage is the formation of the cyclopean image as it involves performing a multiscale Gabor decomposition followed by the computation of weights. The next most expensive stage is the estimation of model parameters (from UGGD/log-normal models) for the construction of the feature vectors. The computation of MSCN coefficients and uncertainty maps incurs some computational expense as well. We believe that Chen et al. [63]'s method is the computationally least expensive of the three methods being compared.

In Su et al. [13], the two most computationally expensive stages are the formation of the CCI and the extraction of BGGD features from the steerable pyramid decomposition of the CCI. We would like to note that the CCI formation is more expensive than the cyclopean image formation in Chen et al. [63] since both the left and right disparity maps are considered. The next most expensive stages are the formation of MSCN coefficients, estimation of the asymmetric GGD and UGGD model parameters in the spatial and multiscale domains respectively, and the estimation of the model parameters of the exponential model for the correlation between subband coefficients. Based on our estimation, we believe that this method is computationally the most expensive of the three.

Finally, in our proposed method, the most computationally intensive stage

is the estimation of the BGGD model parameters at 3 scales and 6 orientations. The other stages including correlation computation and feature consolidation are of significantly lower computational complexity compared to the model fitting stage. We thus conclude that our method is placed in between the two state-of-the-art NR stereo IQA algorithms.

Table 1: SROCC ON LIVE PHASE-I DATASET.

Algorithm	**WN**	**JP2K**	**JPEG**	**BLUR**	**FF**	**ALL**
Benoit [17]	0.930	0.910	0.603	0.931	0.699	0.899
You [19]	0.940	0.860	0.439	0.882	0.588	0.878
Gorley [16]	0.741	0.015	0.569	0.750	0.366	0.142
Chen [18]	0.948	0.888	0.530	0.925	0.707	0.916
Hewage [78]	0.940	0.856	0.500	0.690	0.545	0.814
Akther [38]	0.914	0.866	0.675	0.555	0.640	0.383
Chen [63]	0.919	0.863	0.617	0.878	0.652	0.891
StereoQUE	**0.910**	**0.917**	**0.782**	**0.865**	**0.666**	**0.911**

Table 2: LCC ON LIVE PHASE-I DATASET.

Algorithm	**WN**	**JP2K**	**JPEG**	**BLUR**	**FF**	**ALL**
Benoit [17]	0.925	0.939	0.640	0.948	0.747	0.902
You [19]	0.941	0.877	0.487	0.919	0.730	0.881
Gorley [16]	0.796	0.485	0.312	0.852	0.364	0.451
Chen [18]	0.942	0.912	0.603	0.942	0.776	0.917
Hewage [78]	0.895	0.904	0.530	0.798	0.669	0.830
Akther [38]	0.904	0.905	0.729	0.617	0.503	0.626
Chen [63]	0.917	0.907	0.695	0.917	0.735	0.895
StereoQUE	**0.919**	**0.938**	**0.806**	**0.881**	**0.758**	**0.917**

Table 3: RMSE ON LIVE PHASE-I DATASET.

Algorithm	**WN**	**JP2K**	**JPEG**	**BLUR**	**FF**	**ALL**
Benoit [17]	6.307	4.426	5.022	4.571	8.257	7.061
You [19]	5.621	6.206	5.709	5.679	8.492	7.746
Gorley [16]	10.197	11.323	6.211	7.562	11.569	14.635
Chen [18]	5.581	5.320	5.216	4.822	7.387	6.533
Hewage [78]	7.405	5.530	5.543	8.748	9.226	9.139
Akther [38]	7.092	5.483	4.273	11.387	9.332	14.827
Chen [63]	6.433	5.402	4.523	5.898	8.322	7.247
StereoQUE	**6.664**	**4.943**	**4.391**	**6.938**	**9.317**	**6.598**

Table 4: SROCC ON LIVE PHASE-II DATASET.

Algorithm	**WN**	**JP2K**	**JPEG**	**BLUR**	**FF**	**ALL**
Benoit [17]	0.923	0.751	0.867	0.455	0.773	0.728
You [19]	0.909	0.894	0.795	0.813	0.891	0.786
Gorley [16]	0.875	0.110	0.027	0.770	0.601	0.146
Chen [18]	0.940	0.814	0.843	0.908	0.884	0.889
Hewage [78]	0.880	0.598	0.736	0.028	0.684	0.501
Akther [38]	0.714	0.724	0.649	0.682	0.559	0.543
Chen [63]	0.950	0.867	0.867	0.900	0.933	0.880
S3DBLINQ Index [13]	0.946	0.845	0.818	0.903	0.899	0.905
StereoQUE	**0.932**	**0.864**	**0.839**	**0.846**	**0.860**	**0.888**

Table 5: LCC ON LIVE PHASE-II DATASET.

Algorithm	**WN**	**JP2K**	**JPEG**	**BLUR**	**FF**	**ALL**
Benoit [17]	0.926	0.784	0.853	0.535	0.807	0.748
You [19]	0.912	0.905	0.830	0.784	0.915	0.800
Gorley [16]	0.874	0.372	0.322	0.934	0.706	0.515
Chen [18]	0.957	0.834	0.862	0.963	0.901	0.900
Hewage [78]	0.891	0.664	0.734	0.450	0.746	0.558
Akther [38]	0.722	0.776	0.786	0.795	0.674	0.568
Chen [63]	0.947	0.899	0.901	0.941	0.932	0.895
S3DBLINQ Index [13]	0.953	0.847	0.888	0.968	0.944	0.913
StereoQUE	**0.920**	**0.867**	**0.829**	**0.878**	**0.836**	**0.845**

Table 6: RMSE ON LIVE PHASE-II DATASET.

Algorithm	**WN**	**JP2K**	**JPEG**	**BLUR**	**FF**	**ALL**
Benoit [17]	4.028	6.096	3.787	11.763	6.894	7.490
You [19]	4.396	4.186	4.086	8.649	4.649	6.772
Gorley [16]	5.202	9.113	6.940	4.988	8.155	9.675
Chen [18]	3.368	5.562	3.865	3.747	4.966	4.987
Hewage [78]	10.713	7.343	4.976	12.436	7.667	9.364
Akther [38]	7.416	6.189	4.535	8.450	8.505	9.294
Chen [63]	3.513	4.298	3.342	4.725	4.180	5.102
S3DBLINQ Index [13]	3.547	5.482	4.169	4.453	4.199	4.657
StereoQUE	**4.325**	**5.087**	**4.756**	**6.662**	**6.519**	**7.279**

Table 7: SROCC ON SYMMETRIC AND ASYMMETRIC DISTORTIONS IN LIVE PHASE-II DATASET.

Algorithm	**SYMM**	**ASYMM**
Benoit [17]	0.860	0.671
You [19]	0.914	0.701
Gorley [16]	0.383	0.056
Chen [18]	0.923	0.842
Hewage [78]	0.656	0.496
Akther [38]	0.420	0.517
Chen [63]	0.918	0.834
S3D-BLINQ Index [13]	0.937	0.849
StereoQUE	**0.857**	**0.872**

Table 8: SROCC, LCC ON MICT DATABASE.

Algorithm	**LCC**	**SROCC**
Benoit [17]	0.910	0.902
You [19]	0.864	0.857
Chen [18]	0.864	0.862
Akther [38]	0.765	0.785
Chen [63]	0.913	0.904
S3D-BLINQ Index [13]	0.933	0.917
StereoQUE	**0.935**	**0.936**

All the aforementioned results were obtained using a three scale and six orientation pyramid decomposition. This choice of scales and orientations were determined empirically as it gave the best performance over all the databases used for evaluation. We studied the effect of varying the number of scales and orientations on performance and present it next. The LIVE Phase II database is chosen as a representative database for this study. Fig. 8 shows the variation in performance (LCC, SROCC) as the number of scales are varied at a fixed number of orientations. As the number of scales of decomposition exceed 3, the performance drops. This could be explained by the fewer samples available for parameter estimation at coarser scales. This in turn is related to the image size. Based on this observation, the number of scales was chosen to be 3. Similarly, Fig. 9 shows performance variation as a function of the number of orientations used (at 3 scales). We see that performance variation is very small for 4 or more orientations with the best performance occurring with 6 orientations. We thus chose 6 orientations in our work.

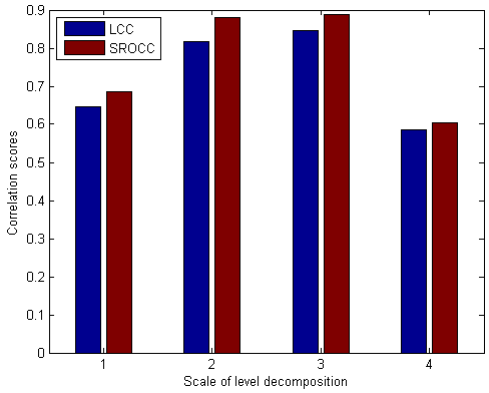


Figure 8: Performance of StereoQUE as a function of the number of scales.

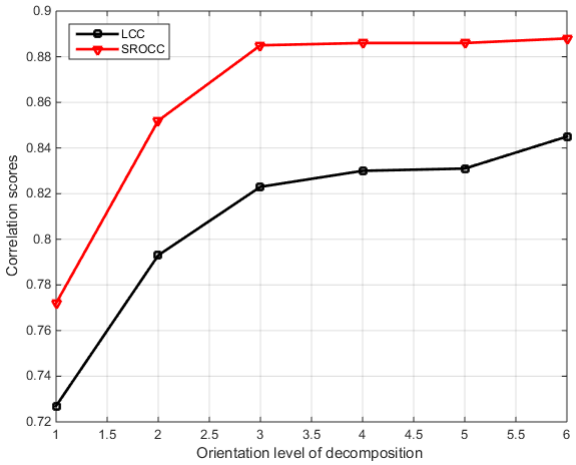


Figure 9: Performance of StereoQUE as a function of the number of orientations.

5. CONCLUSIONS AND FUTURE WORK

We presented a BGGD model for natural stereo scene statistics and an NR stereo IQA algorithm based on the model. The proposed BGGD model accurately captures the joint statistics of luminance and disparity subband coefficients. We believe that this model could be useful in several applications like stereo correspondence, denoising, quality assessment etc.

The utility of the BGGD model was demonstrated in a NR stereo IQA application dubbed StereoQUE. StereoQUE was evaluated on popular stereo image databases and shown to perform competitively with state-of-the-art methods. Further, it delivers consistently high performance on asymmetric distortions

thereby demonstrating the effectiveness of the approach. In future, we intend to demonstrate the utility of the proposed BGGD model in some of the other applications mentioned above.

References

[1] C. Fehn, 3d tv broadcasting, Fraunhofer Institute for Telecommunications, Berlin, Germany, Information Society Technologies, Proposal No. IST-2001-34396.

[2] M. Drake, L. Safonova, R. Tenniswood, '3d's effect on a film's box office and profitability', http://falsecreekproductions.com/wp-content/uploads/2011/01/3Ds-Effect-on-Box-Off-3-14-2011.pdf (2011).

[3] Motion Picture Association of America, Theatrical market statistics 2013, http://www.mpaa.org/wp-content/uploads/2014/03/MPAA-Theatrical-Market-Statistics-2013_032514-v2.pdf (2013).

[4] J. Wang, Z. Wang, Perceptual quality of asymmetrically distorted stereoscopic images: the role of image distortion types, in: Proc. International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM 2014), pp. 29–31.

[5] M.-J. Chen, A. Bovik, L. Cormack, Study on distortion conspicuity in stereoscopically viewed 3d images, in: IVMSP Workshop, 2011 IEEE 10th, 2011, pp. 24–29. doi:10.1109/IVMSPW.2011.5970349.

[6] M. T. Lambooij, W. A. IJsselsteijn, I. Heynderickx, Visual discomfort in stereoscopic displays: a review, in: Electronic Imaging 2007, International Society for Optics and Photonics, 2007, pp. 64900I–64900I.

[7] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, H.-P. Seidel, A perceptual model for disparity, in: ACM Transactions on Graphics (TOG), Vol. 30, ACM, 2011, p. 96.

[8] R. Soundararajan, A. Bovik, Rred indices: Reduced reference entropic differencing for image quality assessment, Image Processing, IEEE Transactions on 21 (2) (2012) 517 –526. doi:10.1109/TIP.2011.2166082.

[9] H. Sheikh, A. Bovik, L. Cormack, No-reference quality assessment using natural scene statistics: Jpeg2000, Image Processing, IEEE Transactions on 14 (11) (2005) 1918 –1927. doi:10.1109/TIP.2005.854492.

[10] A. Moorthy, A. Bovik, Blind image quality assessment: From natural scene statistics to perceptual quality, Image Processing, IEEE Transactions on 20 (12) (2011) 3350 –3364. doi:10.1109/TIP.2011.2147325.

[11] M. A. Saad, A. C. Bovik, C. Charrier, Blind image quality assessment: A natural scene statistics approach in the dct domain, Image Processing, IEEE Transactions on 21 (8) (2012) 3339–3352.

[12] A. Mittal, R. Soundararajan, A. C. Bovik, Making a completely blind image quality analyzer, Signal Processing Letters, IEEE 20 (3) (2013) 209–212.

[13] C.-C. Su, L. K. Cormack, A. C. Bovik, Oriented correlation models of distorted natural images with application to natural stereopair quality evaluation, Image Processing, IEEE Transactions on 24 (5) (2015) 1685–1699.

[14] C.-C. Su, L. K. Cormack, A. C. Bovik, Bivariate statistical modeling of color and range in natural scenes, in: IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics, 2014, pp. 90141G–90141G.

[15] P. Campisi, P. Le Callet, E. Marini, Stereoscopic images quality assessment, in: Proceedings of 15th European Signal Processing Conference (EUSIPCO07), 2007.

[16] P. Gorley, N. Holliman, Stereoscopic image quality metrics and compression, in: Electronic Imaging 2008, International Society for Optics and Photonics, 2008, pp. 680305–680305.

[17] A. Benoit, P. Le Callet, P. Campisi, R. Cousseau, Quality assessment of stereoscopic images, EURASIP journal on image and video processing 2008 (2008) Article–ID.

[18] M.-J. Chen, C.-C. Su, D.-K. Kwon, L. K. Cormack, A. C. Bovik, Full-reference quality assessment of stereopairs accounting for rivalry, Signal Processing: Image Communication 28 (9) (2013) 1143–1155.

[19] J. You, L. Xing, A. Perkis, X. Wang, Perceptual quality assessment for stereoscopic images based on 2d image quality metrics and disparity analysis, in: Proc. of International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, AZ, USA, 2010.

[20] W. Chen, F. Jérôme, M. Barkowsky, P. Le Callet, Exploration of quality of experience of stereoscopic images: Binocular depth, in: Sixth International Workshop on Video Processing and Quality Metrics for Consumer Electronics-VPQM 2012, 2012, pp. 1–6.

[21] R. Bensalma, M.-C. Larabi, Towards a perceptual quality metric for color stereo images, in: Image Processing (ICIP), 2010 17th IEEE International Conference on, IEEE, 2010, pp. 4037–4040.

[22] R. Bensalma, C. Iarabi, A stereoscopic quality metric based on binocular perception, in: Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on, IEEE, 2010, pp. 41–44.

[23] R. Bensalma, M.-C. Larabi, A perceptual metric for stereoscopic image quality assessment based on the binocular energy, Multidimensional Systems and Signal Processing 24 (2) (2013) 281–316.

[24] L. Lei, C. M. Schor, The spatial properties of binocular suppression zone, Vision research 34 (7) (1994) 937–947.

[25] S. Ryu, D. H. Kim, K. Sohn, Stereoscopic image quality metric based on binocular perception model, in: Image Processing (ICIP), 2012 19th IEEE International Conference on, IEEE, 2012, pp. 609–612.

[26] W. Hachicha, A. Beghdadi, F. A. Cheikh, Stereo image quality assessment using a binocular just noticeable difference model, in: Image Processing (ICIP), 2013 20th IEEE International Conference on, IEEE, 2013, pp. 113–117.

[27] Y. Zhang, D. Chandler, 3d-mad: A full reference stereoscopic image quality estimator based on binocular lightness and contrast perception, Image Processing, IEEE Transactions on 24 (11) (2015) 3810–3825. doi:10.1109/TIP.2015.2456414.

[28] K. Li, S. Wang, Q. Jiang, F. Shao, Objective quality assessment for stereoscopic images based on structure-texture decomposition., WSEAS Transactions on Computers 13.

[29] L. Shen, J. Yang, Z. Zhang, Quality assessment of stereo images with stereo vision, in: Image and Signal Processing, 2009. CISP'09. 2nd International Congress on, IEEE, 2009, pp. 1–4.

[30] Y.-H. Lin, J.-L. Wu, Quality assessment of stereoscopic 3d image compression by binocular integration behaviors, Image Processing, IEEE Transactions on 23 (4) (2014) 1527–1542.

[31] C. Hu, F. Shao, G. Jiang, M. Yu, F. Li, Z. Peng, Quality assessment for stereoscopic images by distortion separation, Journal of Software 9 (1) (2014) 37–43.

[32] C. Galkandage, J. Calic, V. De Silva, S. Dogan, A full-reference stereoscopic image quality metric based on binocular energy and regression analysis, in: 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2015, IEEE, 2015, pp. 1–5.

[33] J. V. de Miranda Cardoso, C. D. M. Regis, M. S. de Alencar, Disparity weighting applied to full-reference and no reference stereoscopic image quality assessment, in: Consumer Electronics (ICCE), 2015 IEEE International Conference on, IEEE, 2015, pp. 477–480.

[34] W. Zhou, G. Jiang, M. Yu, Z. Wang, Z. Peng, F. Shao, Reduced reference stereoscopic image quality assessment using digital watermarking, Computers & Electrical Engineering 40 (8) (2014) 104–116.

[35] A. Maalouf, M.-C. Larabi, Cyclop: A stereo color image quality assessment metric, in: Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, IEEE, 2011, pp. 1161–1164.

[36] A. Chetouani, Full reference image quality metric for stereo images based on cyclopean image computation and neural fusion, in: Visual Communications and Image Processing Conference, 2014 IEEE, IEEE, 2014, pp. 109–112.

[37] Z. P. Sazzad, S. Yamanaka, Y. Kawayokeita, Y. Horita, Stereoscopic image quality prediction, in: Quality of Multimedia Experience, 2009. QoMEx 2009. International Workshop on, IEEE, 2009, pp. 180–185.

[38] R. Akhter, Z. P. Sazzad, Y. Horita, J. Baltes, No-reference stereoscopic image quality assessment, in: IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics, 2010, pp. 75240T–75240T.

[39] S. Ryu, K. Sohn, No-reference quality assessment for stereoscopic images based on binocular quality perception, Circuits and Systems for Video Technology, IEEE Transactions on 24 (4) (2014) 591–602.

[40] F. Shao, S. Gu, G. Jang, M. Yu, A novel no-reference stereoscopic image quality assessment method, in: Photonics and Optoelectronics (SOPO), 2012 Symposium on, IEEE, 2012, pp. 1–4.

[41] S. A. Fezza, M.-C. Larabi, No-reference perceptual blur metric for stereoscopic images, in: 3D Imaging (IC3D), 2014 International Conference on, IEEE, 2014, pp. 1–8.

[42] F. Shao, W. Lin, S. Wang, G. Jiang, M. Yu, Blind image quality assessment for stereoscopic images using binocular guided quality lookup and visual codebook, Broadcasting, IEEE Transactions on 61 (2) (2015) 154–165. doi:10.1109/TBC.2015.2402491.

[43] K. Gu, G. Zhai, X. Yang, W. Zhang, A new no-reference stereoscopic image quality assessment based on ocular dominance theory and degree of parallax, in: Pattern Recognition (ICPR), 2012 21st International Conference on, 2012, pp. 206–209.

[44] K. Gu, G. Zhai, X. Yang, W. Zhang, No-reference stereoscopic iqa approach: From nonlinear effect to parallax compensation, Journal of Electrical and Computer Engineering 2012.

[45] Q. Jiang, F. Duan, F. Shao, 3d visual attention for stereoscopic image quality assessment, Journal of Software 9 (7) (2014) 1841–1847.

[46] M. Solh, G. AlRegib, A no-reference quality measure for dibr-based 3d videos, in: Multimedia and Expo (ICME), 2011 IEEE International Conference on, 2011, pp. 1–6. doi:10.1109/ICME.2011.6012169.

[47] E. Y. Lam, J. W. Goodman, A mathematical analysis of the dct coefficient distributions for images, Image Processing, IEEE Transactions on 9 (10) (2000) 1661–1666.

[48] M. J. Wainwright, E. P. Simoncelli, Scale mixtures of gaussians and the statistics of natural images., in: NIPS, Citeseer, 1999, pp. 855–861.

[49] A. K. Moorthy, A. C. Bovik, Blind image quality assessment: From natural scene statistics to perceptual quality, Image Processing, IEEE Transactions on 20 (12) (2011) 3350–3364.

[50] A. Mittal, A. K. Moorthy, A. C. Bovik, Blind/referenceless image spatial quality evaluator, in: Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on, IEEE, 2011, pp. 723–727.

[51] A. Mittal, A. K. Moorthy, A. C. Bovik, No-reference image quality assessment in the spatial domain, Image Processing, IEEE Transactions on 21 (12) (2012) 4695–4708.

[52] H. R. Sheikh, A. C. Bovik, L. Cormack, No-reference quality assessment using natural scene statistics: Jpeg2000, Image Processing, IEEE Transactions on 14 (11) (2005) 1918–1927.

[53] A. Mittal, A. K. Moorthy, J. Ghosh, A. C. Bovik, Algorithmic assessment of 3d quality of experience for images and videos, in: Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE), 2011 IEEE, IEEE, 2011, pp. 338–343.

[54] J. Huang, A. Lee, D. Mumford, Statistics of range images, in: Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, Vol. 1, 2000, pp. 324–331 vol.1. doi:10.1109/CVPR.2000.855836.

[55] O. Wulf, B. Wagner, Fast 3d scanning methods for laser measurement systems, in: International conference on control systems and computer science (CSCS14), 2003, pp. 2–5.

[56] Y. Liu, A. C. Bovik, L. K. Cormack, Disparity statistics in natural scenes, Journal of Vision 8 (11) (2008) 19.

[57] B. Potetz, T. S. Lee, Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes, JOSA A 20 (7) (2003) 1292–1303.

[58] Y. Liu, L. Cormack, A. Bovik, Statistical modeling of 3-d natural scenes with application to bayesian stereopsis, Image Processing, IEEE Transactions on 20 (9) (2011) 2515 –2530. doi:10.1109/TIP.2011.2118223.

[59] Y. Liu, L. K. Cormack, A. C. Bovik, Luminance, disparity, and range statistics in 3d natural scenes, in: IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics, 2009, pp. 72401G–72401G.

[60] K. Kokkinakis, A. K. Nandi, Exponent parameter estimation for generalized gaussian probability density functions with application to speech modeling, Signal Processing 85 (9) (2005) 1852–1858.

[61] Dominguez-Molina, J Armando and González-Farías, Graciela and Rodríguez-Dagnino, Ramó, A practical procedure to estimate the shape parameter in the generalized gaussian distribution, http://www.cimat.mx/reportes/enlinea/I-01-18_eng.pdf (2001).

[62] S. Khan Md, B. Appina, S. Channappayya, Full-reference stereo image quality assessment using natural stereo scene statistics, Signal Processing Letters, IEEE 22 (11) (2015) 1985–1989. doi:10.1109/LSP.2015.2449878.

[63] M.-J. Chen, L. K. Cormack, A. C. Bovik, No-reference quality assessment of natural stereopairs, Image Processing, IEEE Transactions on 22 (9) (2013) 3379–3391.

[64] C.-C. Su, A. C. Bovik, L. K. Cormack, Natural scene statistics of color and range, in: Image Processing (ICIP), 2011 18th IEEE International Conference on, IEEE, 2011, pp. 257–260.

[65] E. P. Simoncelli, W. T. Freeman, The steerable pyramid: A flexible architecture for multi-scale derivative computation, in: Image Processing, International Conference on, Vol. 3, IEEE Computer Society, 1995, pp. 3444–3444.

[66] A. Karasaridis, E. Simoncelli, A filter design technique for steerable pyramid image transforms, in: Acoustics, Speech, and Signal Processing, IEEE International Conference on, Vol. 4, IEEE, 1996, pp. 2387–2390.

[67] http://www.cns.nyu.edu/ eero/steerpyr/.

[68] E. P. Simoncelli, E. H. Adelson, Subband transforms, in: J. W. Woods (Ed.), Subband Image Coding, Kluwer Academic Publishers, Norwell, MA, 1990, Ch. 4, pp. 143–192.

[69] D. Field, What is the goal of sensory coding?, Neural computation 6 (4) (1994) 559–601.

[70] D. Scharstein, R. Szeliski, A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, International journal of computer vision 47 (1-3) (2002) 7–42.

[71] W. J. Levelt, On binocular rivalry, Vol. 2, Mouton The Hague, 1968.

[72] R. Rifkin, A. Klautau, In defense of one-vs-all classification, The Journal of Machine Learning Research 5 (2004) 101–141.

[73] N. Cristianini, J. Shawe-Taylor, An introduction to support vector machines and other kernel-based learning methods, Cambridge university press, 2000.

[74] B. Schölkopf, K.-K. Sung, C. J. Burges, F. Girosi, P. Niyogi, T. Poggio, V. Vapnik, Comparing support vector machines with gaussian kernels to radial basis function classifiers, Signal Processing, IEEE Transactions on 45 (11) (1997) 2758–2765.

[75] C.-C. Chang, C.-J. Lin, Libsvm: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology (TIST) 2 (3) (2011) 27.

[76] A. K. Moorthy, C.-C. Su, A. Mittal, A. C. Bovik, Subjective evaluation of stereoscopic image quality, Signal Processing: Image Communication 28 (8) (2013) 870–883.

[77] Vqeg. (aug. 2003). final report from the video quality experts group on the validation of objective models of video quality assessment, phase ii. [online]. available:
http://www.its.bldrdoc.gov/vqeg/projects/frtv-phase-ii/frtv-phase-ii.aspx.

[78] C. T. Hewage, M. G. Martini, Reduced-reference quality metric for 3d depth map transmission, in: 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2010, IEEE, 2010, pp. 1–4.